Advances in Mathematics: Scientific Journal **9** (2020), no.9, 7429–7438
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.9.89

PERFORMANCE OF THE DOUBLE SAMPLING NP CHART BASED ON THE MEDIAN RUN LENGTH

MOI HUA TUH[1], MING HA LEE, MAN F. LAU, AND PATRICK H. H. THEN

ABSTRACT. **Objectives:** The purpose of this study is to use the median run length (MRL) as a criterion to measure the performance of the double sampling (DS) np chart. **Methods/statistical analysis:** An optimization model of the DS np chart is developed in this study by minimizing the out-of-control MRL. **Findings:** The numerical and graphical results show that the shape of the run length distribution of the DS np chart changes in accordance to the magnitude of shift in the process fraction nonconforming, from highly skewed when there is no process shift to almost symmetric when the process shift is large. The existing DS np chart is evaluated based on the average run length (ARL) criterion. However, the ARL does not provide a clear picture about the chart's performance when the run length distribution is skewed. Herein, we recommend an optimal design of the MRL-based DS np chart which delivers a better interpretation for the performance of the DS np chart. **Application/Improvements:** The optimal parameters of the MRL-based DS np chart are provided in this study.

1. INTRODUCTION

The double sampling (DS) method proposed by [1] was claimed as economically better than the single sampling method. Research on the DS method was gaining interest among researchers as some of the properties of the DS chart are superior to the competing charts [2]. The DS np chart was proposed by [3] and

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 62P30, 49M37.

Key words and phrases. average run length, double sampling np chart, median run length.

the DS np chart was compared with the single-sampling np chart, the variable sample size np chart, the CUSUM np chart and the EWMA np chart. The comparison showed that the DS np chart is the quickest control chart to detect the increasing shift in the process fraction nonconforming.

The average run length (ARL) has been widely used as the performance measure of the control charts over the past few decades. The ARL is the average number of samples plotted on the control chart before an out-of-control signal is triggered. However, many researchers questioned the sole dependence on the ARL as a criterion to measure the chart's performance, such as [4–6]. On the other hand, the median run length (MRL) is a better criterion to interpret the chart's performance compared to the ARL when the run length distribution is skewed (see [4, 5]), where the MRL is the 50th percentile of the run length distribution. As pointed out by [6], the percentiles of the run length distribution including the MRL provide a better indication of the chart's performance. Noted by [5], the use of the MRL develops quality practitioners' confidence to have a better understanding of the performance of a control chart. In view of this advantage, the MRL is proposed as the performance measure to design the control charts by researchers such as [7, 8], to name a few.

The DS np chart proposed by [3] was evaluated based on the ARL criterion. However, the ARL is the average of the run length (RL) that has been reported to be inappropriate in representing "half of the time" [9] and hence quality practitioners should refrain from taking this approach. Consequently, the MRL-based DS np chart is proposed in this study, where the MRL is suggested as an alternative criterion to evaluate the performance of the DS np chart. The main contributions of this study are: (1) to provide a procedure to design the MRL-based DS np chart for the process fraction nonconforming, and (2) to provide the optimal parameters for the MRL-based DS np chart.

2. A Review of the DS np Chart

2.1. The DS np Chart.

The DS np chart proposed by [3] was employed to monitor the quality characteristic of a process, which is the number of nonconforming items in a sample and it is assumed that the underlying process follows a binomial distribution.

The operation procedure of the DS np chart is outlined as follows:

Step 1. The chart limits are set as $WL = Ac_1 + 0.5$, $CL_1 = Re - 0.5$ and $CL_2 = Ac_2 + 0.5$, where WL, CL_1 and CL_2 are the warning limit for the first stage of the DS scheme, the control limit for the first stage of the DS scheme and the control limit for the second stage of the DS scheme, respectively. In these expressions, Ac_1, Re_1 and Ac_2 are the acceptance number in the first sample, the rejection number in the first sample and the acceptance number in the second stage, respectively.

Step 2. A sample of size n_1 is taken and the number of nonconforming items d_1 in this first sample is counted, in which the procedure is at the first stage of the DS scheme.

Step 3. (a) If $d_1 < WL$, the process is considered as in- control and return to Step 2.

(b) If $d_1 > CL_1$, the process is concluded to be out-of-control and a corrective action is taken to search and remove the assignable cause(s), then return to Step 2.

(c) If $WL < d_1 < CL_1$, the process is considered as in-control and the procedure goes to the second stage of the DS scheme, in which a second sample of size n_2 is taken and the number of nonconforming items d_2 in this second sample is counted. After that, proceed to the next step.

Step 4. If $(d_1 + d_2) < CL_2$, the process is concluded to be in-control and return to Step 2. Otherwise, the process is concluded as out-of-control and a corrective action is taken to search and remove the assignable cause(s), then return to Step 2.

Let $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the round down and round up to the nearest integer, respectively. The probability of a plotted sample falling within the chart limits of the DS np chart [3] is given by

$$A = A_1 + A_2, \tag{2.1}$$

where

$$\begin{aligned} A_1 &= P\left(d_1 \leq \lfloor WL \rfloor\right) \\ &= \sum_{d_1=0}^{\lfloor WL \rfloor} C_{d_1}^{n_1} p^{d_1} \left(1-p\right)^{n_1-d_1} \end{aligned} \tag{2.2}$$

and

$$\begin{aligned} A_2 &= P\left(\lfloor WL \rfloor < d_1 < \lceil CL_1 \rceil\right) \cap P\left(d_1 + d_2 \leq \lfloor CL_2 \rfloor\right) \\ &= \sum_{d_1 = \lfloor WL \rfloor + 1}^{\lceil CL_1 \rceil - 1} \left[C_{d_1}^{n_1} p^{d_1} \left(1-p\right)^{n_1 - d_1} \left(\sum_{d_2 = 0}^{\lfloor CL_2 \rfloor - d_1} C_{d_2}^{n_2} p^{d_2} \left(1-p\right)^{n_2 - d_2} \right) \right]. \end{aligned} \tag{2.3}$$

Here, $C_d^n = \frac{n!}{d!(n-d)!}$. Note that A_1 is the probability that $d_1 < WL$ at the first stage of the DS scheme, while A_2 is the probability that $WL < d_1 < CL_1$ at the first stage of the DS scheme and $(d_1 + d_2) < CL_2$ at the second stage of the DS scheme.

The efficiency of the DS np chart is determined by its speed in detecting an increasing shift in the process fraction nonconforming p with the magnitude of shift $\delta = p_1/p_0$, where $p_1 > p_0$. Note that $p = p_0$ when $\delta = 1$, while $p = p_1$ when $\delta > 1$. Here, p_0 is the in-control fraction nonconforming and p_1 is the out-of-control fraction nonconforming.

2.2. **The Run Length Properties.** Since the RL of the control chart is geometric distributed, the probability mass function (pmf) and the cumulative distribution function (cdf) of the RL in general [10] are defined as

$$f_{RL}(l) = (1 - A)A^{l-1} \tag{2.4}$$

and

$$F_{RL}(l) = P(RL \leq l) = 1 - A^l, \tag{2.5}$$

respectively, for $l = 1, 2, 3, \ldots$ Using the cdf in (2.5), [6] presented a simplified method, in which the 100αth $(0 < \alpha < 1)$ percentile of the RL can be computed as the smallest integer τ such that

$$\tau \geq \frac{\ln(1 - \alpha)}{\ln A}. \tag{2.6}$$

This enables the percentile value to be computed easily. Since the MRL is the smallest RL with a cumulative probability of at least 50% of the time, then $\alpha = 0.5$ in (2.6). Note that $MRL = MRL_0$ is the in-control MRL when $\delta = 1$, whereas $MRL = MR_1$ is the out-of-control MRL when $\delta > 1$.

As suggested by [3], the average sample size of the DS np chart is computed as

$$ASS = n_1 + n_2 A_r, \tag{2.7}$$

where $A_r = P\left(\lfloor WL \rfloor < d_1 < \lceil CL_1 \rceil\right)$ is the probability of taking the second sample of size n_2. Two ASS values are usually of interest, namely the in-control ASS (ASS_0) and the out-of-control ASS (ASS_1).

3. Optimal Design of the DS np Chart Based on the MRL

An optimization model of the DS np chart based on the MRL is developed in this section by minimizing the objective function which is expressed as

$$\min_{n_1, n_2, WL, CL_1, CL_2} MRL_1\left(\delta_{opt}\right) \tag{3.1}$$

subject to

$$MRL_0 \geq MRL_{0\,\min}, \tag{3.2}$$

$$ASS_0 = n, \tag{3.3}$$

$$1 \leq n_1 < n < n_1 + n_2 \text{ and } n_1 < n_2, \tag{3.4}$$

where $MRL_1(\delta_{opt})$ is the MRL_1 value at the desired shift size δ_{opt} in the process fraction nonconforming p for a fast detection, $MRL_{0\,\min}$ is the minimum desired MRL_0 and n is the desired ASS_0.

A control chart is considered superior to its competitors if this control chart gives the smallest MRL_1 value. Similarly, when designing the DS np chart at a desired shift, the parameters that produce the lowest MRL_1 are identified as the optimal parameters from all the possible $(n_1, n_2, WL, CL_1, CL_2)$ combinations. The optimization procedure is detailed below:

Step 1. Specify the desired values of $MRL_{0\,\min}$, δ_{opt}, p_0 and n.

Step 2. All possible pairs of (n_1, n_2) are selected based on the constraints in (3.4). For each pair of (n_1, n_2), the CL_1 value is determined for any fixed value of WL, then the CL_2 value is determined for any fixed values of WL and CL_1. Here, the values of WL, CL_1 and CL_2 are adjusted such that $MRL_0 \geq MRL_{0\,\min}$ and $ASS_0 = n$. The values of WL, CL_1 and CL_2 are determined based on Step 1 of the operation procedure discussed in previous section, where $0 < WL < CL_1 \leq CL_2$. Therefore, we can obtain all the possible $(n_1, n_2, WL, CL_1, CL_2)$ combinations that satisfy the constraints in (3.2) – (3.4).

Step 3. For each possible $(n_1, n_2, WL, CL_1, CL_2)$ combination in Step 2, compute the MRL_1 by means of (2.6) (when $p = p_1$ and $\alpha = 0.5$) and search for the optimal $(n_1, n_2, WL, CL_1, CL_2)$ combination with the lowest $MRL_1(\delta_{opt})$ value. Since the MRL is an integer, it can be happened that more than one optimal $(n_1, n_2, WL, CL_1, CL_2)$ combinations that produce the similar lowest $MRL_1(\delta_{opt})$ value. For such a case, the optimal $(n_1, n_2, WL, CL_1, CL_2)$ combination that delivers the lowest ASS_1 value is selected as the optimal parameter combination of the DS np chart.

An optimization MATLAB program is written by incorporating the aforementioned three-step procedure to compute the optimal $(n_1, n_2, WL, CL_1, CL_2)$ combination for the DS np chart.

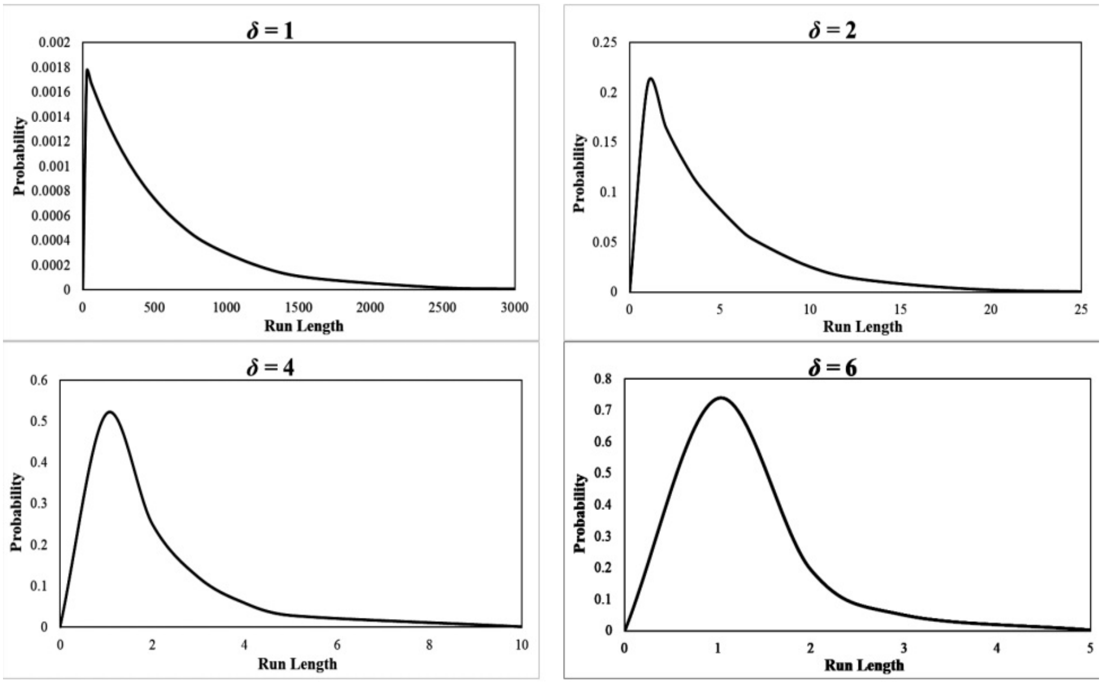

FIGURE 1. Plots of the pmf of the RL for the DS np chart with $(n_1, n_2, WL, CL_1, CL_2) = (43, 2276, 1.5, 5.5, 34.5)$ when $\delta_{opt} = 1.5$, $p_0 = 0.01$, $n = 200$ and $MRL_{0\,\min} = 370.4$ for various shift sizes in p

4. Performance of the DS np Chart Based on the MRL

The plots of the pmf of the RL for the DS np chart in Figure 1 are obtained by considering $(n_1, n_2, WL, CL_1, CL_2) = (43, 2276, 1.5, 5.5, 34.5)$ for $\delta_{opt} = 1.5$, $n = 200$, $p0 = 0.01$ and $MRL_{0\,\min} = 370.4$. As demonstrated in Figure 1, the skewness of the RL distribution of the DS np chart changes from highly right skewed when there is no shift ($\delta = 1$) to almost symmetric for the large shift size.

In this study, the DS np chart is optimally designed based on the MRL, where the MRL values are computed by assuming that the underlying process follows a binomial distribution. The optimal $(n_1, n_2, WL, CL_1, CL_2)$ combination of the DS np chart as well as the corresponding MRL_0 and ARL_0 values are presented in Table 1 with $MRL_{0\,\min} \in \{200, 370.4\}$, $\delta_{opt} \in \{1.5, 2, 3\}$, $p_0 \in \{0.005, 0.01, 0.02\}$ and $n \in \{25, 50, 100, 200, 400, 800\}$. These optimal $(n_1, n_2, WL, CL_1, CL_2)$ combinations will facilitate quality practitioners to choose the optimal parameters of the MRL-based DS np chart in the implementation of the DS np chart. For example, when $\delta_{opt} = 2$, $MRL_{0\,\min} = 200$, $p_0 = 0.02$ and $n = 50$, the optimal $(n_1, n_2, WL, CL_1, CL_2)$ combination is $(25, 282, 1.5, 4.5, 12.5)$. In addition, Table 1 provides a comparison between the ARL_0 and MRL_0 values of the DS np chart. It is noted that all the MRL_0 values are smaller than the corresponding ARL_0 values.

5. The Performance of the DS np Chart Based on the Percentiles of the Run Length Distribution

The investigation of the percentiles of the RL distribution is useful in providing an in-depth understanding about the entire RL distribution of a control chart including the early false alarm rates. In this study, the percentiles of the RL distribution is calculated using (2.6). Table 2 displays the ARL and some of the percentiles of the RL distribution for the DS np chart with $(n_1, n_2, WL, CL_1, CL_2) = (43, 2276, 1.5, 5.5, 34.5)$ when $\delta_{opt} = 1.5$, $n = 200$, $p_0 = 0.01$ and $MRL_{0\,\min} = 370.4$.

The lower percentiles such as the 5th, 10th and 20th percentiles of the RL distribution for the in-control condition ($\delta = 1$) provide information about the early false alarm rates. For example, the 10th percentile of the in-control RL

TABLE 1. Optimal parameters $(n_1, n_2, WL, CL_1, CL_2)$ with the corresponding MRL_0 and ARL_0 for the DS np chart

δ_{opt}	p_0	n	$MRL_{0\min} = 200$							$MRL_{0\min} = 370.4$						
			n_1	n_2	WL	CL_1	CL_2	MRL_0	ARL_0	n_1	n_2	WL	CL_1	CL_2	MRL_0	ARL_0
1.5	0.005	100	8	2340	0.5	2.5	17.5	205	294.82	47	2285	1.5	3.5	18.5	375	541.15
		200	74	2366	1.5	4.5	19.5	212	305.13	63	3430	1.5	5.5	26.5	374	539.64
		400	103	3151	1.5	4.5	25.5	200	288.05	174	3928	2.5	6.5	31.5	371	534.33
		800	133	4650	1.5	6.5	35.5	200	288.80	320	6127	3.5	11.5	46.5	371	534.51
	0.01	50	4	1167	0.5	2.5	17.5	209	301.69	25	970	1.5	4.5	16.5	376	541.73
		100	37	1192	1.5	4.5	19.5	204	294.29	35	1358	1.5	5.5	22.5	371	534.92
		200	50	1677	1.5	5.5	26.5	200	288.33	43	2276	1.5	5.5	34.5	372	536.09
		400	68	2238	1.5	6.5	34.5	200	288.52	159	3153	3.5	8.5	47.5	372	535.82
	0.02	25	2	580	0.5	2.5	17.5	221	318.03	11	719	1.5	3.5	21.5	371	535.00
		50	18	646	1.5	4.5	20.5	201	290.36	16	852	1.5	5.5	26.5	412	593.55
		100	24	921	1.5	4.5	28.5	218	313.94	41	1214	2.5	7.5	36.5	373	538.18
		200	32	1254	1.5	5.5	37.5	206	296.28	106	1512	4.5	9.5	46.5	376	541.67
2	0.005	100	23	708	0.5	2.5	8.5	200	288.00	13	1379	0.5	2.5	13.5	435	627.36
		200	100	1115	1.5	4.5	12.5	224	323.27	104	1002	1.5	4.5	12.5	376	542.84
		400	226	1655	2.5	6.5	17.5	210	302.92	50	1578	0.5	4.5	16.5	378	545.09
		800	535	1995	4.5	10.5	22.5	208	300.42	768	769	7.5	12.5	16.5	375	541.05
	0.01	50	34	352	1.5	4.5	8.5	202	291.78	8	543	0.5	3.5	11.5	379	546.93
		100	50	559	1.5	4.5	12.5	226	325.44	13	710	0.5	3.5	14.5	381	549.81
		200	116	756	2.5	7.5	16.5	200	288.60	25	787	0.5	4.5	16.5	394	568.76
		400	268	994	4.5	9.5	22.5	200	288.96	384	388	7.5	12.5	16.5	372	536.62
	0.02	25	5	208	0.5	2.5	9.5	264	381.12	4	270	0.5	2.5	11.5	398	573.29
		50	25	282	1.5	4.5	12.5	224	323.19	26	253	1.5	4.5	12.5	387	558.17
		100	54	494	2.5	5.5	19.5	201	289.68	9	547	0.5	3.5	20.5	467	674.10
		200	111	486	3.5	9.5	21.5	202	290.70	195	297	8.5	11.5	19.5	434	626.06
3	0.005	100	84	240	1.5	4.5	5.5	215	310.22	77	406	1.5	3.5	7.5	374	539.79
		200	184	252	2.5	4.5	7.5	203	293.03	182	281	2.5	6.5	7.5	379	546.63
		400	278	301	1.5	6.5	8.5	220	316.84	381	438	4.5	7.5	11.5	490	707.18
		800	534	534	2.5	8.5	13.5	263	379.19	425	426	0.5	7.5	12.5	374	539.19
	0.01	50	42	120	1.5	4.5	5.5	220	317.61	16	228	0.5	3.5	7.5	453	652.88
		100	92	127	2.5	4.5	7.5	209	300.76	91	141	2.5	6.5	7.5	386	556.87
		200	139	150	1.5	6.5	8.5	229	329.80	180	186	3.5	7.5	10.5	450	648.66
		400	237	238	1.5	7.5	13.5	221	318.30	212	213	0.5	7.5	12.5	393	567.22
	0.02	25	21	61	1.5	4.5	5.5	221	318.83	8	114	0.5	2.5	7.5	379	546.42
		50	46	64	2.5	4.5	7.5	222	319.85	37	77	1.5	5.5	7.5	384	554.33
		100	69	77	1.5	6.5	8.5	230	331.92	64	98	1.5	6.5	9.5	383	552.02
		200	96	121	0.5	6.5	13.5	209	301.60	105	108	0.5	7.5	12.5	443	638.64

distribution is 57, indicating that there is a probability of 0.1 that a false alarm will occur by the 57th sample. On the other hand, it is about half of the time that a false alarm will occur by the 372nd sample, indicating that there is a probability of 0.5 that a false alarm will be detected by sample 372, while the ARL_0 is given as 536.09.

The higher percentiles of the RL distribution for the out-of-control condition ($\delta > 1$) provide information about the out-of-control that will be issued by the control chart at a certain magnitude of shift with a high probability. For example,

TABLE 2. ARLs and percentiles of the RL distribution for the DS np chart with $(n_1, n_2, WL, CL_1, CL_2) = (43, 2276, 1.5, 5.5, 34.5)$ when $\delta_{opt} = 1.5$, $n = 200$, $p_0 = 0.01$ and $MRL_{0\,\min} = 370.4$

δ	ARL	Percentiles of the run length distribution												
		1st	5th	10th	20th	30th	40th	50th	60th	70th	80th	90th	95th	99th
1.0	536.09	6	28	57	120	192	274	372	491	645	862	1234	1605	2467
1.1	161.29	2	9	17	36	58	83	112	148	194	259	371	482	741
1.2	63.39	1	4	7	15	23	33	44	58	76	102	145	189	290
1.3	30.91	1	2	4	7	11	16	22	28	37	49	71	92	141
1.4	17.93	1	1	2	4	7	9	13	16	21	29	41	53	81
1.5	11.93	1	1	2	3	5	6	8	11	14	19	27	35	53
2.0	4.80	1	1	1	1	2	3	3	4	6	7	10	13	20
3.0	2.69	1	1	1	1	1	2	2	2	3	4	5	7	10
4.0	1.93	1	1	1	1	1	1	1	2	2	3	4	5	7
5.0	1.56	1	1	1	1	1	1	1	1	2	2	3	3	5

the 90th percentile is 27 at $\delta = 1.5$, indicating that the DS np chart signals within the first 27 samples with a probability of 0.9.

It can be observed that the difference between the values of ARL_0 and MRL_0 is large when there is no shift ($\delta = 1$) but this difference diminishes as the shift increases. This shows that the RL distribution can be highly skewed when there is no shift and the shape of the RL distribution changes with the magnitude of shift δ.

6. Conclusions

In this study, the implementation of the MRL as a performance measure for the DS np chart has been studied and discussed. The optimal DS np chart is defined as the control chart with the smallest MRL_1 value for a specified shift in the process fraction nonconforming, the desired ASS_0 value and the desired MRL_0 value. The MRL is a more suitable performance measure for the DS np chart compared to the ARL as the shape of the RL distribution changes in accordance to the shift in the process fraction nonconforming. The RL properties of the DS np chart are also furnished via the percentiles of the RL distribution in this study.

References

[1] R. Croasdale: *Control charts for a double-sampling scheme based on average production run lengths*, Int. J. Prod. Res., **12**(5) (1974), 585–592.

[2] J.J. Daudin: *Double sampling $\bar{X}$ charts*, J. Qual. Technol., **24**(2) (1992), 78–87.

[3] DE A.A.A. RODRIGUES, E.K. EPPRECHT, M.S. DE MAGALHÃES: *Double-sampling control charts for attributes*, J. Appl. Stat., **38**(1) (2011), 87–112.

[4] F.F. GAN: *An optimal design of ewma control charts based on median run length*, J. Stat. Comput. Simul., **45**(3–4) (1993), 169–184.

[5] F.F. GAN: *An optimal design of cumulative sum control chart based on median run length*, Commun. Stat. - Simul. Comput., **23**(2) (1994), 485–503.

[6] S. CHAKRABORTI: *Run length distribution and percentiles: The Shewhart $\bar{X}$ chart with unknown parameters*, Qual. Eng., **19**(2) (2007), 119–127.

[7] W.L. TEOH, M.B.C. KHOO, P. CASTAGLIOLA, S. CHAKRABORTI: *Optimal design of the double sampling $\bar{X}$ chart with estimated parameters based on median run length*, Comput. Ind. Eng., **67** (2014), 104–115.

[8] M. H. LEE, M.B.C. KHOO: *Optimal design of synthetic np control chart based on median run length*, Commun. Stat. - Theory Methods, **46**(17) (2017), 8544–8556.

[9] A.C. PALM: *Tables of run length percentiles for determining the sensitivity of Shewhart control charts for averages with supplementary runs rules*, J. Qual. Technol., **22**(4) (1990), 289–298.

[10] D. BROOK, D.A. EVANS: *An approach to the probability distribution of cusum run length*, *Biometrika*, **59**(3) (1972), 539–549.

FACULTY OF COMPUTER AND MATHEMATICAL SCIENCES,
UNIVERSITI TEKNOLOGI MARA,
CAWANGAN SARAWAK, KOTA SAMARAHAN, SARAWAK, MALAYSIA
Email address: tuhmoihua@uitm.edu.my

FACULTY OF ENGINEERING, COMPUTING AND SCIENCE
SWINBURNE UNIVERSITY OF TECHNOLOGY SARAWAK CAMPUS
KUCHING, SARAWAK, MALAYSIA
Email address: mhlee@swinburne.edu.my

SCHOOL OF SOFTWARE AND ELECTRICAL ENGINEERING
SWINBURNE UNIVERSITY OF TECHNOLOGY
MELBOURNE, AUSTRALIA
Email address: elau@swin.edu.au

FACULTY OF ENGINEERING, COMPUTING AND SCIENCE
SWINBURNE UNIVERSITY OF TECHNOLOGY SARAWAK CAMPUS
KUCHING, SARAWAK, MALAYSIA
Email address: pthen@swinburne.edu.my